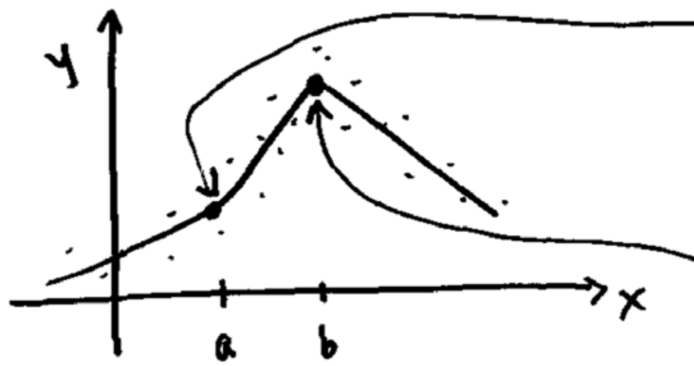
Linear Spline model



$$\beta_0+\beta_1 a = \beta_0+\delta_0+(\beta_1+\delta_1)a$$
$$0=\delta_0+\delta_1 a$$
$$\delta_0 = -\delta_1 a$$

$$\beta_0+\delta_0+(\beta_1+\delta_1)b =$$
$$\beta_0+\delta_0+\gamma_0+(\beta_1+\delta_1+\gamma_1)b$$
$$0=\gamma_0+\gamma_1 b$$
$$\gamma_0 = -\gamma_1 b$$

Segment 1: $y = \beta_0+\beta_1 x+\varepsilon$

Segment 2: $y = (\beta_0+\delta_0)+(\beta_1+\delta_1)x+\varepsilon$

Segment 3: $y = (\beta_0+\delta_0+\gamma_0)+(\beta_1+\delta_1+\gamma_1)x+\varepsilon$

Quadratic splines:

each segment looks like $y = \beta_0+\beta_1 x+\beta_2 x^2+\varepsilon$

Continuity: use restrictions to make the y-values equal at the endpoints

Differentiability: make the 1st derivatives match

Cubic splines

$$y = \beta_0+\beta_1 x+\beta_2 x^2+\beta_3 x^3+\varepsilon$$

Continuity, differentiability, plus make the 2nd derivatives match

Stat 564 HW#6

5.1 Byers and Williams ("Viscosities of Binary and Ternary Mixtures of Polyaromatic Hydrocarbons," *Journal of Chemical and Engineering Data*, **32**, 349–354, 1987) studied the impact of temperature (the regressor) on the viscosity (the response) of toluene-tetralin blends. The following table gives the data for blends with a 0.4 molar fraction of toluene.

Temperature (°C)	Viscosity (mPa · s)
24.9	1.133
35.0	0.9772
44.9	0.8532
55.1	0.7550
65.2	0.6723
75.2	0.6021
85.2	0.5420
95.2	0.5074

a. Plot a scatter diagram. Does it seem likely that a straight-line model will be adequate?

b. Fit the straight-line model. Compute the summary statistics and the residual plots. What are your conclusions regarding model adequacy?

c. Basic principles of physical chemistry suggest that the viscosity is an exponential function of the temperature. Repeat part b using the appropriate transformation based on this information.

5.13 Schubert et al. ("The Catapult Problem: Enhanced Engineering Modeling Using Experimental Design," *Quality Engineering*, **4**, 463–473, 1992) conducted an experiment with a catapult to determine the effects of hook (x_1), arm length (x_2), start angle (x_3), and stop angle (x_4) on the distance that the catapult throws a ball. They threw the ball three times for each setting of the factors. The following table summarizes the experimental results.

x_1	x_2	x_3	x_4		y	
–1	–1	–1	–1	28.0	27.1	26.2
–1	–1	1	1	46.3	43.5	46.5
–1	1	–1	1	21.9	21.0	20.1
–1	1	1	–1	52.9	53.7	52.0
1	–1	–1	1	75.0	73.1	74.3
1	–1	1	–1	127.7	126.9	128.7
1	1	–1	–1	86.2	86.5	87.0
1	1	1	1	195.0	195.9	195.7

a. Fit a first-order regression model to the data and conduct the residual analysis.

b. Use the sample variances as the basis for weighted least-squares estimation of the original data (not the sample means).